Supporting information

Estrogen Receptor α/β – Cofactor Motif Interactions; Interplay of Tyrosine 537/488 Phosphorylation and LXXLL Motifs

Hoang D. Nguyen,[a,b,c] Trang T.P. Phan,[a,b,c] Maelle Carraz,[a,c] and Luc Brunsveld*[a,c]

[a] *Laboratory of Chemical Biology, Department of Biomedical Engineering, Technische Universiteit Eindhoven, Den Dolech 2, Eindhoven, the Netherlands. Fax +31 40 2478367; Tel: +31 40 2473737; E-mail: l.brunsveld@tue.nl.*
[b] *Center for Bioscience and Biotechnology and Laboratory of Molecular Biotechnology, University of Science, Vietnam National University Ho Chi Minh City, 227 Nguyen Van Cu, District 5, Ho Chi Minh, Vietnam.*
[c] *Chemical Genomics Centre of the Max Planck Society, Otto-Hahn Strasse 15, D-44227, Dortmund, Germany.*

Contents Page

Supporting Figures Page

Supporting Tables Page

Table S1. Peptides and proteins used in this study.

Name of peptide or protein	Method	Mass calculated (Da)	Mass determined (Da)
SRC1Box2	Chemical synthesis	2557	2557
Fluorescein-SRC1Box2	Chemical synthesis	2962	2962
P1	Chemical synthesis	2128	2128
P2	Chemical synthesis	2059	2059
P4	Chemical synthesis	2122	2122
P5	Chemical synthesis	2140	2140
ERα-H12 (ERα 529-553)	Chemical synthesis	2762	2762
ERα-H12-pY537	Chemical synthesis	2842	2842
ERα LBD	Protein expression	32402	32402
Thioester-ERα-LBD	Protein expression	29362	29362
ERα LBD-W	Semi-synthesis	31984	31984
ERα LBD-P	Semi-synthesis	32063	32066
ERβ H12 (ERβ 481-502)	Chemical synthesis	2497	2497
ERβ H12-pY488	Chemical synthesis	2577	2577
ERβ LBD	Protein expression	31042	31042
Thioester-ERβ-LBD	Protein expression	28266	28266
ERβ LBD-W	Sem-isynthesis	30622	30622
ERβ LBD-P	Sem-isynthesis	30701	30701

Samples were measured by either MALDI-TOF or LC-ESI-MS. Protein sequences without methionine at the N terminus were calculated.

1. Production of reference proteins

Reference proteins, ERα LBD (pHT503) [1], ERβ LBD (pHT504) [2], ECFP (pHT490), SRC1Box2- ECFP (pHT508) and ECFP-SRC1Box2 (pHT510) were constructed, expressed and purified as follows.

For construction of plasmids carrying fusion genes encoding for His-SNAP-ECFP-His (renamed ECFP), SRC1Box2-SNAP-ECFP-His (renamed SRC1Box2-ECFP) and His-SNAP-ECFP-SRC1Box2 (renamed ECFP-SRC1Box2), plasmids pHT490, pHT508 and pHT510 were constructed as follows. SNAP-ECFP-encoding fusion gene was amplified using primers ON005 and ON010 (Table S2) and pHT406 as template (described in another publication). The *Sac*I/*Xho*I treated PCR fragments were ligated with pTriEx4-Neo/*Sac*I/*Xho*I resulting pHT490. To construct plasmid carrying a fusion with coactivator SRC1Box2 encoding gene, the DNA encoding for the SRC1Box2 was generated by the hybridization of two complement oligonucleotides ON161 and ON162 (table S2), and this DNA fragment was then cloned into plasmid pHT406 at *Bam*HI and *Nco*I resulting pHT508. Similarly, the DNA fragment of two complement ON121 and ON122 (Table S2) was cloned into plasmid pHT490 at *Xho*I resulting pHT510.

For expression and purification of reference proteins, plasmids pHT503 [1], pHT504 [2], pHT490, pHT508 and pHT510 were transformed into *E. coli* Rosetta 2 (DE3) pLacI cells (Novagen). A single colony was then grown in LB medium with 100 mg/L ampicilline and 34 mg/L chloramphenicol at 37 °C. At OD_{600} nm ~0.8, protein expression was induced with 0.5 mM IPTG for 16–18 hr at 16 °C. Bacteria were collected by centrifugation at 5000 rpm for 10 min. The cells were resuspended in washing buffer (50 mM sodium phosphate, 300 mM NaCl, 5% glycerol and 20 mM imidazole (pH 8.0)) containing 1 mM PMSF and disrupted by a sonicator (Branson) and centrifuged to collect the clear supernatant. The sample was loaded on a HisTrap HP 5 ml column (GE Healthcare) and the protein-bound column was washed with 100 ml washing buffer. Proteins were eluted by imidazole and fractions containing the Strep-His-ERα-LBD (renamed ERα LBD, Figure S1 for the mass measurement, List S2 for protein sequence) and Strep-His-ERβ-LBD (renamed ERβ LBD, Figure S2 for the mass measurement, List S2), ECFP, SRC1Box2-ECFP and ECFP-SRC1Box2 (List S2) proteins were desalted using on a PD-10 column (GE Healthcare) pre-equilibrated with storage buffer (25 mM NaCl, 20 mM Tris-HCl (pH 8.0), 10% glycerol, and 1 mM TCEP).

Table S2. Oligonucleotides used in this study.

Primer	Sequence (5´- 3´)	Description
ON005	ggcggatccggcgagctcgacaaagactgcgaaatgaagcg	For pHT490
ON010	Ggcctcgagcttgtacagctcgtccatgccg	For pHT490
ON047	tatggcaagctggagccacccgcagttcgaaaagtctgc	Strep-tag for pHT505 and pHT506
ON048	tagcagacttttcgaactgcgggtggctccagcttgcca	Strep-tag for pHT505 and pHT506
ON121	tcgacctgaccgctcgtcacaaaatcctgcatcgtctgctgcaggaaggttcttaag	For pHT510
ON122	tcgacttaagaaccttcctgcagcagacgatgcaggattttgtgacgagcggtcagg	For pHT510
ON123B	Caggagctgtcgtattccagtc	Binding on T7 DNA
ON124B	Aacccctcaagacccgtttag	Binding on T7 DNA
ON158	ggccatacatgtctggtaagaagaacagcctggccttgtcc	Binding on ERα for pHT501
ON158B	ttttttgctcttctgcacttcatgctgtacagatgctc	Binding on ERα for pHT501
ON159	ggccatacatgtctggtctggacgccctgagccccga	Binding on ERβ for pHT502
ON159B	ttttttgctcttctgcacttcatgttgagcagatgttccatg	Binding on ERβ for pHT502
ON161	catggaagcgagcctgaccgctcgtcacaaaatcctgcatcgtctgctgcaggaag	For pHT508
ON162	gatccttcctgcagcagacgatgcaggattttgtgacgagcggtcaggctcgcttc	For pHT508
ON165F	gccatgaattctnnynnynnynnynnynnynnyctgnnynnyctgctgnnynnynnynnynnynnynnyaagcttgcggccgcactcgagtaa	For the library
ON166R	Ttactcgagtgcggccgcaagctt	For the library
ON193	Ccacggtggtcttcgcccag	Binding on T7 DNA

N = A, C, G or T; Y = C or T.

Table S3. List of plasmids used in this study.

Name of plasmid	Properties or Name of protein	Reference
pTriEx-Neo	Plasmid containing promoters for *E. coli*, insect and mammalian cells	Novagen
pET15b-ERα-LBD	Carrying a ERα LBD-encoding gene	Gift from Bayer-Schering Pharma
pET15b-ERß-LBD	Carrying a ERß LBD-encoding gene	Gift from Bayer-Schering Pharma
pHT406	Carrying a Strep-SNAP-ECFP-His-encoding gene	Brunsveld Laboratory collection
pHT449	Carrying His-MCS-intein-CBD-encoding gene	[3]
pHT490	ECFP*	This study
pHT501	Carrying a His- ERα LBD-encoding gene	This study
pHT502	Carrying a His-ERß LBD-encoding gene	This study
pHT503	ERα LBD*	[1]
pHT504	ERß LBD*	[2]
pHT505	Thioester-ERα-LBD	This study
pHT506	Thioester-ERß-LBD	This study
pHT508	SRC1Box2-ECFP*	This study
pHT510	ECFP-SRC1Box2*	This study

* Name of the proteins used in this study.

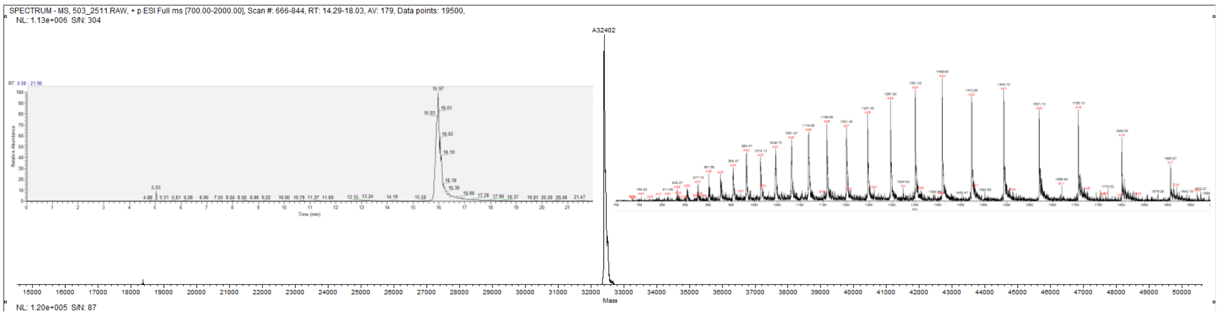

Figure S1. LC-ESI-MS analysis of ERα LBD (mass without methionine at the N terminus, calculated 32402 Da and measured 32402 Da).

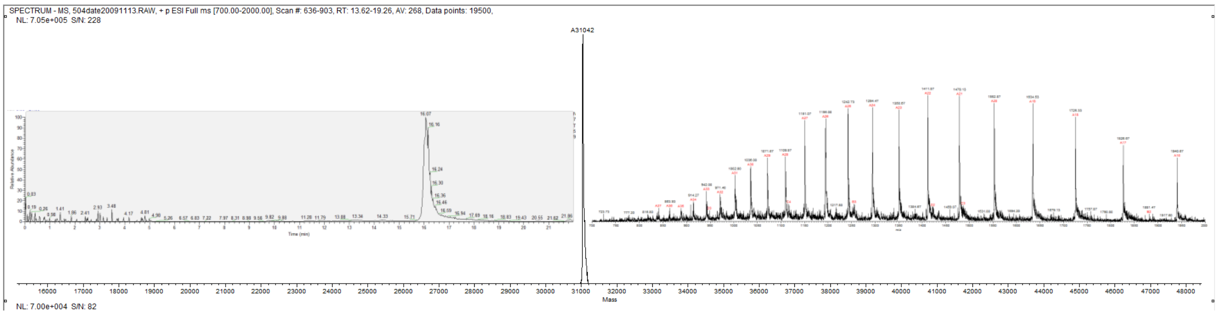

Figure S2. LC-ESI-MS analysis of ERß LBD (mass without methionine at the N terminus, calculated 31042 Da and measured 31042 Da).

2. Production of thioester ERα-LBD and ERβ-LBD

Construction of plasmids pHT505 and pHT506 for expression of thioester ERα-LBD and ERβ-LBD (without H12) was described as follows. The pET15b-ERα-LBD and pET15b-ERβ-LBD plasmids (gifts from Bayer-Schering Pharma AG) encoding the LBD of the human ERα and ERβ were used as template using ON158 and ON158B for ERα and ON159 and ON159B for ERβ. The PCR products were digested by *Afl*III and *Sap*I to generate *Nco*I and *Sap*I compatible ends. The resulting DNA fragments were cloned into pHT449 [3] at *Nco*I and *Sap*I resulting pHT501 and pHT502, respectively. To generate fusions Strep-His-ERα-LBD and Strep-His-ERβ-LBD, the DNA encoding for Strep-tag resulting from the hybridization of two complementary oligonucleotides, ON047 and ON048 was inserted into pHT501 and pHT502 resulting pHT505 and pHT506, respectively.

For expression of thioester ERα-LBD and ERβ-LBD, the resulting plasmids, pHT505 and pHT506 were transformed into *E. coli* Rosetta 2 (DE3) cells. The cells were cultured and expressed in LB medium containing ampicillin (100 mg/L) and chloramphenicol (34 mg/L). When the culture reached to 0.8 at OD600, the cells were induced with 0.5 mM IPTG at 30° C for 8 hours. After harvested, the cell pellet was resuspended and sonicated in lysis buffer (20 mM HEPES pH 8.0, 500 mM NaCl, 10% glycerol, 1 mM EDTA and 10 µM beta-estradiol, 20 µM PMSF, 0.1 mM TCEP and 0.5% Triton X-100). The lysate was centrifuged and the supernatant was load on the columns (EconoPac, Biorad) containing chitin beads (NEB). The beads were washed with 10 column volumes of washing buffer (20 mM HEPES pH 8.0, 500 mM NaCl, 10% glycerol, 1 mM EDTA and 10 µM beta-estradiol). Cleavage of the fusion proteins was induced by suspending the beads in cleavage buffer (20 mM Tris-HCl pH8.5, 10% glycerol, 10µM beta-estradiol and 500 mM MESNA) and overnight incubation at 4 °C. After collecting the thioester ERα-LBD, this protein was concentrated to 300 – 500 µM (10 - 15 mg/mL). The proteins were stored at -80 °C until ligation.

Semisynthesis of nonphosphorylated (ERα LBD-W and ERβ LBD-W) and phosphorylated proteins (ERα LBD-P and ERβ LBD-P) were described [4] with modifications. Components in the ligation reaction include 200 µM thioesther ERα LBD or thioesther ERß LBD, 800 µM peptide, 200 mM MESNA, 20 mM Tris-HCl pH 8.0, 10 mM NaCl, 10% glycerol, 1 mM TCEP. The purity of the resulting proteins was examined by SDS-PAGE (Figure 3Sa), the presence of a phosphorylated tyrosine on ERα LBD-P and ERβ LBD-P were confirmed by Western blot (Figure 3Sb).

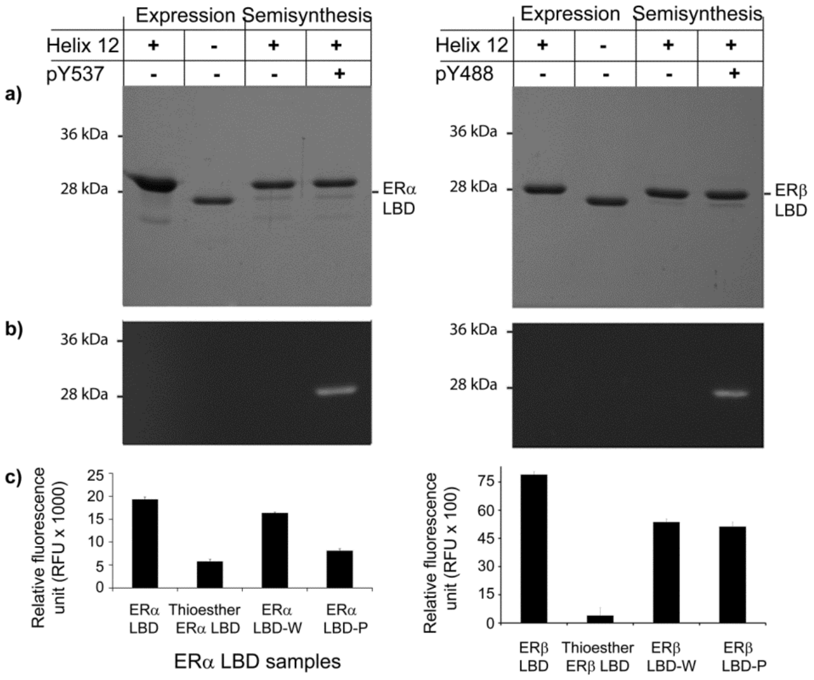

Figure S3. Analysis of semisynthesized ERα and ERβ. (a) SDS-PAGE analysis, (b) Western blot against using monoclonal anti-phosphotyrosine-FITC antibody and (c) functional analysis of immobilized protein on Strep-Tactin-coated plate.

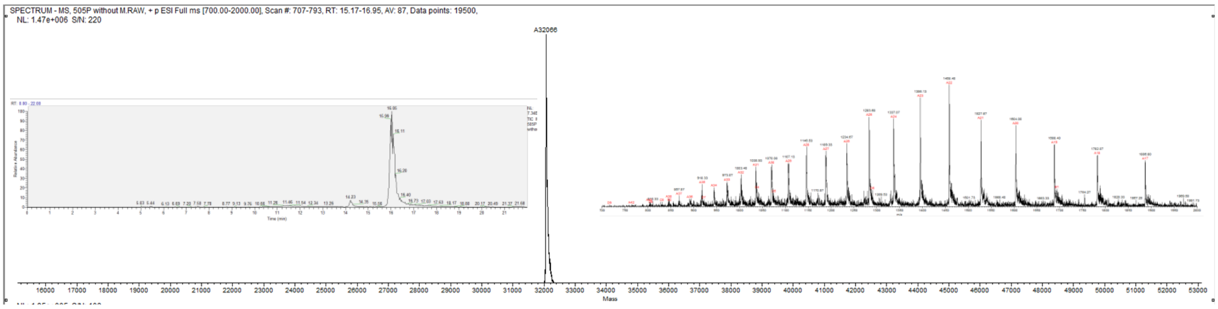

Figure S4. LC-ESI-MS analysis of ERα LBD-P (mass without methionine at the N terminus, calculated 32063 Da and measured 32066 Da).

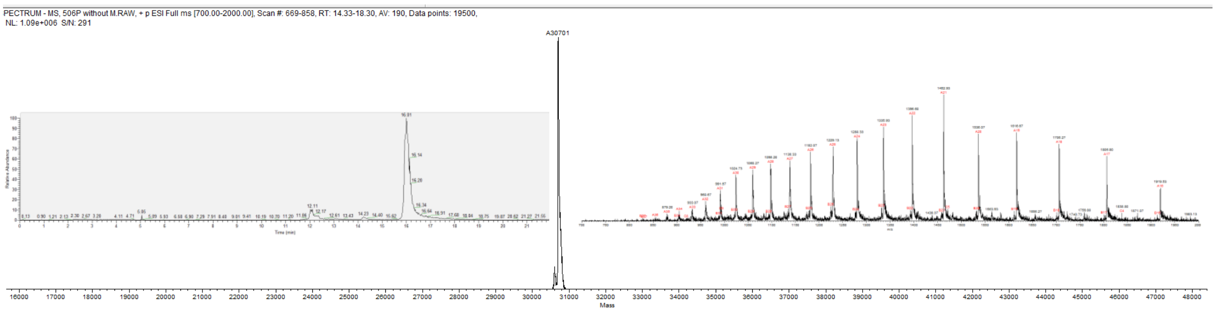

Figure S5. LC-ESI-MS analysis of ERß LBD-P (mass without methionine at the N terminus, calculated 30701 Da and measured 30701 Da).

3. Displacement assays and functional analysis of immobilized ERs

The competitive fluorescence polarization assay was carried out as described by Vaz *et. al* [5].

For functional analysis of immobilized constructs, eight strep-tag fusion proteins were diluted to 10 µM in the binding buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl and 10 µM β-estradiol) containing 0.5% BSA. The above diluted proteins (200 µl) were incubated in Strep-Tactin coated microplates (IBA) at 4 °C for an hour. After incubation, the coated plate was washed 3 times using the binding buffer. As a control surface, similarly block microtiter wells that have not been coated with the target molecules. To test the function of immobilized proteins on the surface, the coated wells were incubated with 200 µl of 1 µM ECFP-SRC1Box2 and SRC1Box2-ECFP with ECFP as a negative control in the binding buffer for 1 hour using TBST buffer (TBS with 0.5% Tween 20). The bound proteins were eluted with either by biotin or 1% SDS and the fluorescence signal was determined. The wells were measured using the plate reader (Safire2, Tecan) with excitation filter 440 ± 15 nm and emission filter 480 ± 15 nm. The Figure S6 showed that the immobilized ERß LBD on the surface could bind to the fusion proteins containing SRC1Box2 (SRC1Box2-ECFP and ECFP-SRC1Box2) while the fusion protein without SRC1Box (ECFP) showed only the background signals. These result indicated that ERß LBD are functionally immobilized on the surface and this method could be used to determine the function of the proteins on the surface. This established method using SDS to elute the protein was used to check for all the other ERs proteins as shown in Figure S3c.

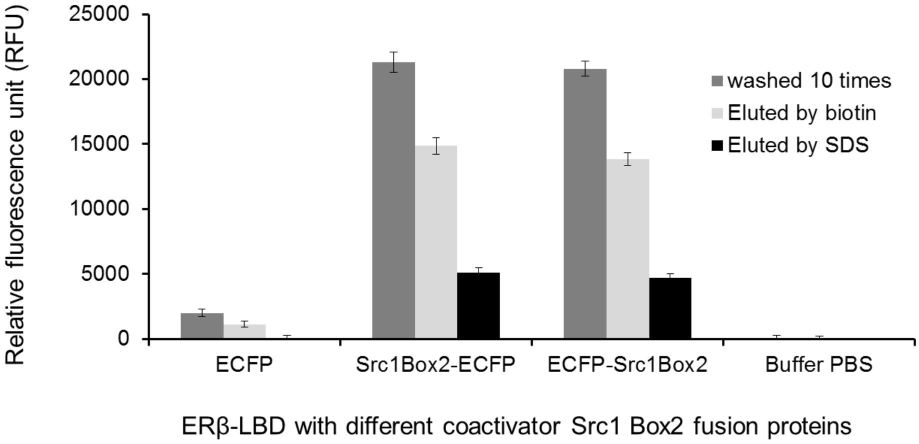

Figure S6. Analysis of immobilized estrogen receptors on Strep-Tacin-coated surface. ERβ LBD in a fusion form with Strep-tag II at the N-terminal was immobilized on the Strep-Tactin-coated wells. Fusion proteins of coactivator SRC1Box2 with ECFP (SRC1Box2-ECFP or ECFP-SRC1Box2) were added to these wells. The interactions of SRC1Box2 with ERβ LBD were measured by fluorescence signal remaining in the well after 10 times washing with PBS buffer. The well could be used to measure fluorescence immediately. In addition, the eluted fluorescence proteins by using biotin and SDS were also measured.

Table S4. Peptide sequences from phage display selections with their pI and hydrophobicity.

ERα LBD

(19 sequences from round 3 and 12 sequences from round 4 (R4))

Name	Sequence	pI of C terminus	Mean hydrophob-icity
27 R4	FVSRSNSLLGLLLDGAINP	3.08	1.54
20 R4	VNPVSSRLIALLYDDVNDG	3.42	1.01
13	HRLLVNSLDVLLSVDGPNL	3.8	1.76
15	SFTSTGLLNILLSATNADS	3.8	0.59
17	FRNGCFNLYRLLDSNFAAN	3.8	0.18
11	PYRHRALLGFLLNDFSDHH	5.05	0.89
7	CRYHSIILSHLLSSGDYSH	5.06	0.26
31 R4	RNFSRNILYNLLHCPLNDN	5.08	-0.24
8	DIFSSSNLSRLLSSPYYPG	5.24	0.43
25 R4	RDSRHCSLSSLLSGGCSVI	5.24	-0.85
1	FASGHSTLRDLLCNGGSYS	5.52	-0.52
5	SSCGTSRLCSLLNYSVVYS	5.52	-1.57
16	ADFPSATLNRLLLCYSNNL	5.52	0.73
26 R4	VGRSHFALGTLLSRDFALC	5.55	0.97
12	VRRYSFHLAALLSSASGGH	6.46	-0.38
22 R4	GDGADRFLFNLLVSGGSIH	6.71	1.25
10	ACDISHSLRYLLLHSCGSS	6.73	0.20
21 R4	LPPASSLLFNLLHSSNINY	6.74	1.38
28 R4	FGIGTSALGHLLSASGCRS	7.96	-0.06
29 R4	GSYNSPNLRCLLVGRDSCR	8.22	-1.57
6	CSDGGSILTLLLCSNVSRA	8.25	-0.23
2	GTDGRCTLRCLLTRPYSNP	8.41	-0.60
4	PCRSSSVLYHLLSSRGYGF	8.46	0.21
9	SAISGSFLYNLLSSRVYSN	8.46	0.19
3	GRVSGSRLRDLLRSYPAHT	8.75	-1.57
19	VLGVHHSLSNLLRGSLYTS	8.75	0.46
23 R4	AVGNNSTLFNLLRNTNHYV	8.75	-0.76
30 R4	INRPGTDLARLLRYSISVS	8.75	-0.18
14	VVVYTNILHNLLSRNNTNF	9.47	0.10
24 R4	RTSNVSRLFSLLSRSSLAF	9.74	-0.25
18	LDRVRGILFNLLGTPFRSV	9.75	1.40

ERα LBD-P

(19 sequences from round 3 and 8 sequences from round 4 (R4))

Name	Sequence	pI of C terminus	Mean Hydrophob-icity
23 R4	DHHTNGSLYFLLSNNGDNN	3.8	-0.85
24 R4	HRLLVNSLDVLLSVDGPNL	3.8	1.76
19	FSVAIGNLGTLLVLDLDRN	4.21	2.07
10	AFNRPSFLFNLLTSDTHSL	5.05	1.06
22 R4	PYRHRALLGFLLNDFSDHH	5.05	0.89
8	FSVNSFRLYHLLATSHGTD	5.08	0.43
11	APSPHVNLISLLTSSSSSG	5.19	-0.35
3	PIHGSSILHNLLSGNFTAF	5.24	1.16
12	FASGHSTLRDLLCNGGSYS	5.52	-0.52
13	SSCGTSRLCSLLNYSVVYS	5.52	-0.60
27 R4	TDFHFCTLSPLLDVTRYSY	5.83	1.26
18	IGCRFHYLSGLLPFGGPTF	5.96	1.90
5	VYCVGSRLYYLLSSNCHPS	6.45	-0.02
7	VRRYSFHLAALLSSASGGH	6.46	-0.38
4	RDARPSALYRLLSRHHDGC	6.66	-1.56
2	SYTSHSTLACLLIGRDVHN	6.74	-0.29
16	YTVHFGDLRCLLDRPFCHS	6.74	0.78
17	CSDGGSILTLLLCSNVSRA	8.25	0.21
15	GTDGRCTLRCLLTRPYSNP	8.41	-1.57
1	PCRSSSVLYHLLSSRGYGF	8.46	-0.23
20 R4	AVTAHSRLYALLGVGYPYR	8.59	0.24
6	GRVSGSRLRDLLRSYPAHT	8.75	-1.57
26 R4	FHDPHGTLFGLLSGRSSFG	9.47	0.84
21 R4	NIPANSVLFNLLSAGVNHR	9.49	0.04
9	GSYNNGALFDLLIPSTRST	9.75	0.04
25 R4	GSYNNGALFDLLIPSTRST	9.75	0.04
14	DTVCFSRLRRLLISSHHNR	9.76	-1.09

Sequences in grey boxes are appeared at least twice. pI of C terminus, pI of the last seven amino acids at the C terminus. Mean hydrophobicity, average hydrophobicity of all amino acid that was calculated by using combined consensus scale [6].

ERβ LBD
(26 sequence from round 3 and additional 12 sequence from round 4 (R4))

Name	Sequence	pI of C terminus	Mean hydrophobi
14	VNPVSSRLIALLYDDVNDG	3.42	1.01
18	YLHTRHFLISLLSDDGPNN	3.56	0.81
35 R4	DCVGARTLVALLNDCDAVV	3.56	0.91
8	NDYSHGRLRHLLNDFCTGF	3.8	-0.17
15	DNVTGIFLHSLLLGNDPVG	3.8	1.81
17	YPDTHLLLYSLLLPDSANA	3.8	2.07
21	SGRIGSALICLLSGPTDVS	3.8	0.51
26	VFARVPSLDILLDFGCGLP	3.8	2.88
27 R4	LDDHHGLLTHLLGDYSYSF	3.8	1.99
30 R4	LTGRRNVLTALLFDGSLVG	3.8	1.08
37 R4	FPASYTTLRSLLVGYADSF	3.83	1.26
6	YFSYSSILDRLLSIDGDYH	4.2	1.77
22	AISRGRVLISLLHDHYDSL	5.05	1.09
29 R4	PYRHRALLGFLLNDFSDHH	5.05	0.89
4	VSSTNNTLRHLLSAAGYYG	5.24	-0.90
19	SRIHGNFLFHLLSGNVGSN	5.24	0.17
25	RAVCFRSLAYLLTDGRSVN	5.5	-0.29
31 R4	RAVCFRSLXYLLTDGRSVN	5.5	-0.23
5	FASGHSTLRDLLCNGGSYS	5.52	-0.52
12	SSCGTSRLCSLLNYSVVYS	5.52	-0.60
13	VSTARLRLSSLLGNSYGLG	5.52	-0.06
34 R4	FASGHSTLRDLLCNGGSYS	5.52	-0.52
7	IIDDHGSLLSLLPLIGTSF	5.96	3.21
28 R4	VRRYSFHLAALLSSASGGH	6.46	-0.38
10	YGRLSYNLVSLLGSVGSHN	6.74	-0.01
16	DDGSHIALHPLLFGHLVPF	6.74	2.71
11	SNHPGVNLRYLLHNVHSNF	6.92	-0.57
20	VALHGNNLSFLLGHSTGHS	6.92	0.13
1	CSDGGSILTLLLCSNVSRA	8.25	0.21
33 R4	CSDGGSILLLLLCSNVSRA	8.25	0.93
36 R4	SHDNRNDLCYLLRSTNVCF	8.25	-0.74
23	GTDGRCTLRCLLTRPYSNP	8.41	-1.57
32 R4	GTDGRCTLRCLLTRPYSNP	8.41	-1.57
2	PCRSSSVLYHLLSSRGYGF	8.46	-0.23
9	GRVSGSRLRDLLRSYPAHT	8.75	-1.57
38 R4	INRPGTDLARLLRYSISVS	8.75	-0.18
3	FVSASSVLFTLLNARSSGV	9.75	0.76
24	NAFLRGFLADLLAFPSNHR	9.8	1.15

ERβ LBD-P
(23 sequences from round 3 and 11 sequences from round 4 (R4))

Name	Sequence	pI of C terminus	Mean hydrophobi
14	VNPVSSRLIALLYDDVNDG	3.42	1.01
12	YSSGSCILHCLLDSNLCDC	3.56	0.65
13	HTCGSSILLGLLNVDCSYD	3.56	1.01
2	VPLFDCYLSVLLDPGLPVN	3.8	3.09
5	CSPFLINLSRLLLIPDFNL	3.8	3.23
15	SDGFGTLLIGLLYDFVGNP	3.8	2.69
33 R4	RGAVGSSLVCLLTYSDNVC	3.8	0.04
34 R4	ALNSNVGLVRLLLRDDFSS	4.21	0.96
26 R4	PYRHRALLGFLLNDFSDHH	5.05	0.89
32 R4	TYNYHPLLVSLLGANFDGH	5.08	1.22
9	PIHGSSILHNLLSGNFTAF	5.24	1.16
4	RVPNLGILGRLLFGGATIP	5.52	1.49
7	APGRVFSLGRLLNSSSFVS	5.52	0.14
29 R4	FASGHSTLRDLLCNGGSYS	5.52	-0.52
20	YLSNYHNLIRLLGRASYDP	5.84	0.31
25 R4	YRPCHLSLISLLNHDSHLN	5.97	0.52
22	RNHISPGLFYLLPYRNIID	6.26	1.38
6	VRRYSFHLAALLSSASGGH	6.46	-0.38
10	RLHDISFLEYLLSGSGHVV	6.46	1.35
17	CFALVSRLFCLLVCTISHA	6.7	2.12
1	CGHVDGLLSALLGYPVPCR	8.22	1.00
11	CSDGGSILTLLLCSNVSRA	8.25	0.21
23	GYAYRSNLNFLLGSNCRNF	8.25	-0.56
24 R4	SSFCFGSLYSLLATHRCSS	8.3	0.01
31 R4	GTDGRCTLRCLLTRPYSNP	8.41	-1.57
8	PCRSSSVLYHLLSSRGYGF	8.46	-0.23
27 R4	SAISGSFLYNLLSSRVYSN	8.46	0.19
18	ICSNSSRLSDLLVSSSYLR	8.72	-0.23
21	GGNRFSPLYRLLVRYGSNS	8.72	-0.85
16	GRVSGSRLRDLLRSYPAHT	8.75	-1.57
30 R4	FDVTSSGLASLLTRVFSFS	9.41	1.35
19	SDFNRAILNRLLVIVRSGT	9.72	0.30
28 R4	FVSASSVLFTLLNARSSGV	9.75	0.76
3	NAFLRGFLADLLAFPSNHR	9.8	1.15

Sequences in grey boxes appeared at least twice. pI of C terminus, pI of the last seven amino acids at the C terminus. Mean hydrophobicity, average hydrophobicity of all amino acids that was calculated by using combined consensus scale [6].

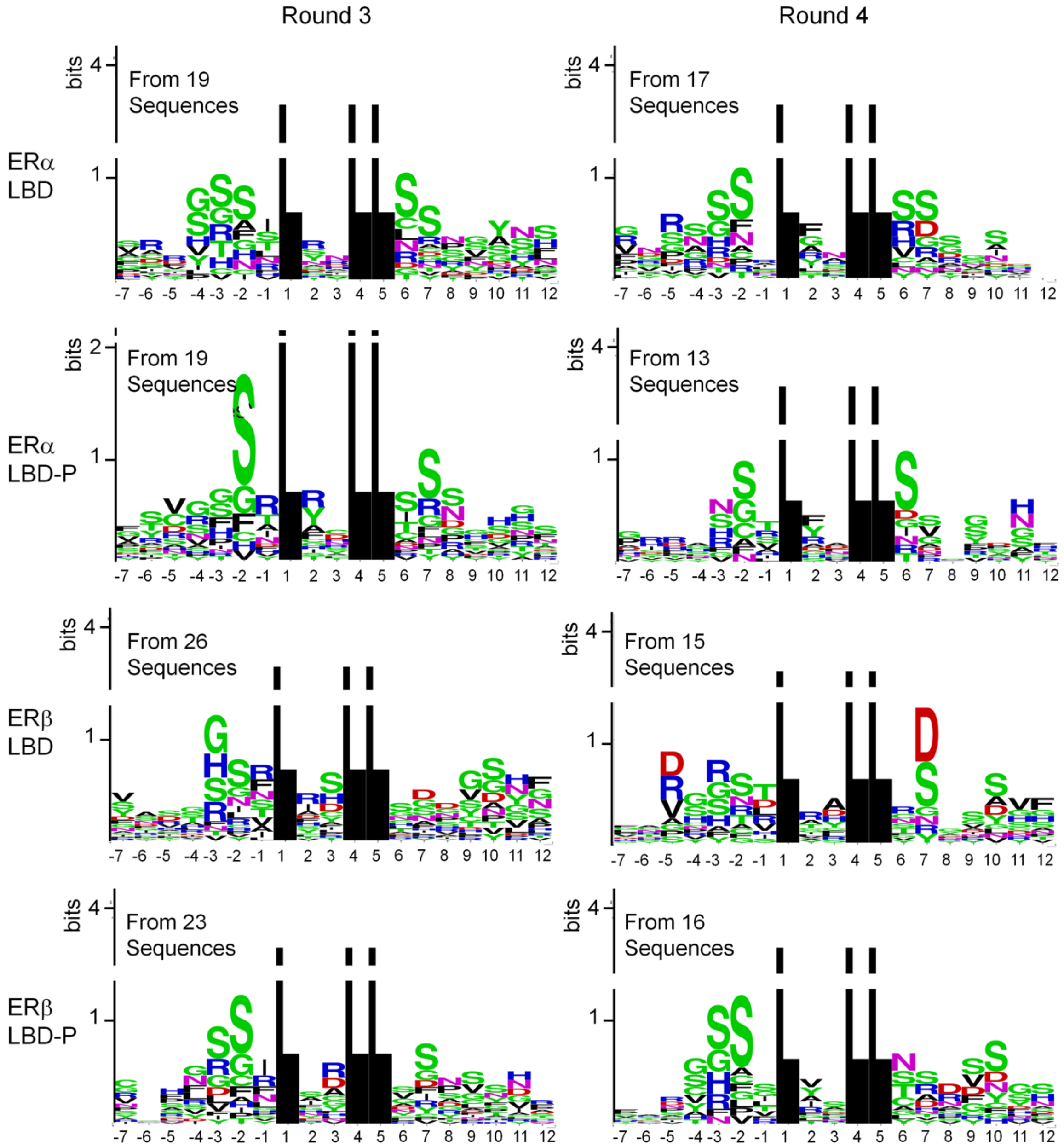

Figure S7. Sequence logo of peptides collected from phage display screening after round three and round four and selection against different ER constructs.

4. Calculation of pI

The pIs were calculated by using the Compute pI/Mw tool (http://web.expasy.org/compute_pi/) [7]. The results were shown in Table S4 and Figure 2.

Table S5. LXXLL pI of C terminus of coactivators, SRC-1, SRC-2 and SRC-3, corepressor RIP140 and collected peptides from M13 phage screening.

	Name	N Terminus	LXXLL motif	C Terminus	pI of C terminus
	SRC1Box1	YSQTSHK	LVQLL	TTTAEQ	4.0
	SRC1Box2	LTERHKI	LHRLL	QEGSPSD	3.67
	SRC1Box3	ESKDHQL	LRYLL	DKDEKDL	4.23
	SRC2Box1	DSKGQTK	LLQLL	TTKSDQM	5.50
	SRC2Box2	LKEKHKI	LHRLL	QDSSSPV	3.80
	SRC2Box3	KKKENAL	LRYLL	DKDDTKD	4.14
	SRC3Box1	LQEKHRI	LHKLL	QNGNSPA	5.52
	SRC3Box2	KKENNAL	LRYLL	DRDDPSD	3.77
	RIP140Box1	DSIVLTY	LEGLL	MHQAAGG	6.49
	RIP140Box2	GKQDSTL	LASLL	QSFSSRL	9.75
	RIP140Box3	YGVASSH	LKTLL	KKSKVKD	10.0
	RIP140Box4	PSVACSQ	LALLL	SSEAHLQ	5.22
	RIP140Box5	QAANNSL	LLHLL	KSQTIPK	10.0
	RIP140Box6	NSHQKVT	LLQLL	LGHKNEE	5.4
	RIP140Box7	LLERRTV	LQLLL	GNPNKGK	10.0
	RIP140Box8	SFSKNGL	LSRLL	RQNQDSY	5.84
	RIP140Box9	ESKSFNV	LKQLL	LSENCVR	5.99
Class I	D2	GSEPKSR	LLELL	SAPVTDV	3.80
	D 11	VESGSSR	LMQLL	MANDLLT	3.80
	D30	HPTHSSR	LWELL	MEATPTM	4.0
Class II	D 14	QEAHGPL	LWNLL	SRSDTDW	4.21
	D 47	HVYQHPL	LLSLL	SSEHESG	4.51
	C 33	HVEMHPL	LMGLL	MESQWGA	4.0
Class III	F 6	GHEPLTL	LERLL	MDDKQAV	4.21
	D 22	LPYEGSL	LLKLL	RAPVEEV	4.53
	D 48	SGWENSI	LYSLL	SDRVSLD	4.21
	D 43	AHGESSL	LAWLL	SGEYSSA	4.0
	D 17	GVFCDSI	LCQLL	AHDNARL	6.79
	D 41	HHNGHSI	LYGLL	AGSDAPS	3.8
	D 26	LGERASL	LDMLL	RQENPAW	6.0
	D 40	SGWNESI	LYRLL	QADAFDV	3.56
	D 15	PSGGSSV	LEYLL	SAPVERE	4.53
	F 4	PVGEPGL	LWRLL	SAPVERE	4.53

Peptides belonging to the three classes were found by [8].

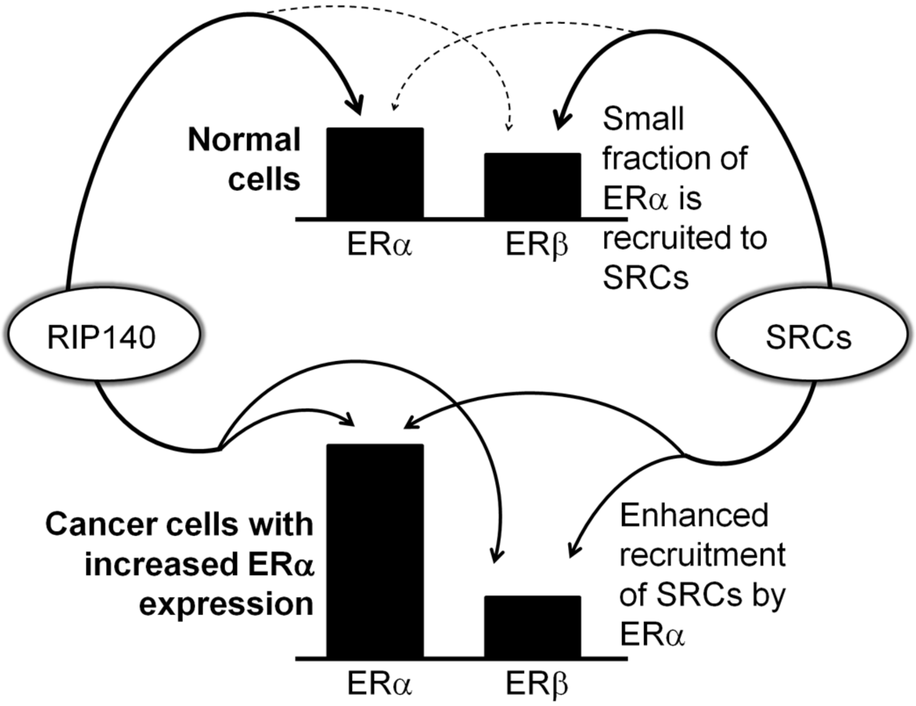

Figure S8. Model for the interplay between ERα overexpression and cofactor recruitment.

Table S6. Protein sequences used in this study.

ERα LBD (recombinant expressed using pHT503 for Strep-His-ERα LBD)
MASWSHPQFEKGSMGSSHHHHHHSSGLVPRGSHMKKNSLALSLTADQMVSALLDAEPPILYSEY
DPTRPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWR
SMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVY
TFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHL
YSMKCKNVVPLYDLLLEMLDAHRLHAPT

ERß LBD (recombinant expressed using pHT504 for Strep-His-ERß LBD)
MASWSHPQFEKGSMGSSHHHHHHSSGLVPRGSHMLDALSPEQLVLTLLEAEPPHVLISRPSAPFT
EASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG
KLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQ
DADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKC
KNVVPVYDLLLEMLNAHVLRG

Thioester-ERα-LBD (recombinant expressed using pHT505 for thioester form of Strep-His-ERα LBD without H12)
MASWSHPQFEKSAMEASHHHHHHHHAMSGKKNSLALSLTADQMVSALLDAEPPILYSEYDPTR
PFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEH
PGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSS
TLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHLYSMK
-thioester

Thioester-ERß-LBD (recombinant expressed using pHT506 for thioester form of Strep-His-ERß LBD without H12)
MASWSHPQFEKSAMEASHHHHHHHHAMSGLDALSPEQLVLTLLEAEPPHVLISRPSAPFTEASM
MMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPGKLIFA
PDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQDADS
SRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMK-
thioester

ERα LBD-P (semisynthesis using thioester-ERα-LBD and ERα-H12); Phosphorylated tyrosine is in bold and underlined.
MASWSHPQFEKGSMGSSHHHHHHSSGLVPRGSHMKKNSLALSLTADQMVSALLDAEPPILYSEY
DPTRPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWR
SMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVY
TFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHL
YSMKCKNVVPL**pY**DLLLEMLDAHRLHAPT

ERß LBD-P (semisynthesis using thioester-ERß-LBD and ERß-H12); Phosphorylated tyrosine is in bold and underlined.
MASWSHPQFEKGSMGSSHHHHHHSSGLVPRGSHMLDALSPEQLVLTLLEAEPPHVLISRPSAPFT
EASMMMSLTKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEVLMMGLMWRSIDHPG
KLIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAMILLNSSMYPLVTATQ
DADSSRKLAHLLNAVTDALVWVIAKSGISSQQQSMRLANLLMLLSHVRHASNKGMEHLLNMKC
KNVVPV**pY**DLLLEMLNAHVLRG

ECFP (recombinant expressed using pHT490 for His-SNAP-ECFP-His)
MAHHHHHHSSGKETAAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKSPGFSSTISELDKDCEM
KRTTLDSPLGKLELSGCEQGLHEIKLLGKGTSAADAVEVPAPAAVLGGPEPLMQATAWLNAYFH

QPEAIEEFPVPALHHPVFQQESFTRQVLWKLLKVVKFGEVISYQQLAALAGNPAATAAVKTALSG
NPVPILIPCHRVVSSSGAVGGYEGGLAVKEWLLAHEGHRLGKPGLGEFVSKGEELFTGVVPILVE
LDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTWGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNY
ISHNVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDP
NEKRDHMVLLEFVTAAGITLGMDELYKLEHHHHHHHH

SRC1Box2-ECFP (recombinant expressed using pHT508 for SRC1Box2-SNAP-ECFP-His); SRC1Box2 is underlined.
MEASLTARHKILHRLLQEGSGELDKDCEMKRTTLDSPLGKLELSGCEQGLHEIKLLGKGTSAADA
VEVPAPAAVLGGPEPLMQATAWLNAYFHQPEAIEEFPVPALHHPVFQQESFTRQVLWKLLKVVK
FGEVISYQQLAALAGNPAATAAVKTALSGNPVPILIPCHRVVSSSGAVGGYEGGLAVKEWLLAH
EGHRLGKPGLGEFVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGK
LPVPWPTLVTTLTWGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFE
GDTLVNRIELKGIDFKEDGNILGHKLEYNYISHNVYITADKQKNGIKANFKIRHNIEDGSVQLADH
YQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKLEHHHHH
HHH

ECFP-SRC1Box2 (recombinant expressed using pHT510 for His-SNAP-ECFP- SRC1Box2); SRC1Box2 is underlined.
MAHHHHHHSSGKETAAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKSPGFSSTISELDKDCEM
KRTTLDSPLGKLELSGCEQGLHEIKLLGKGTSAADAVEVPAPAAVLGGPEPLMQATAWLNAYFH
QPEAIEEFPVPALHHPVFQQESFTRQVLWKLLKVVKFGEVISYQQLAALAGNPAATAAVKTALSG
NPVPILIPCHRVVSSSGAVGGYEGGLAVKEWLLAHEGHRLGKPGLGEFVSKGEELFTGVVPILVE
LDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTWGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNY
ISHNVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDP
NEKRDHMVLLEFVTAAGITLGMDELYKLDLTARHKILHRLLQEGS

5. Reference list

1. M. Carraz, W. Zwart, T. Phan, R. Michalides, and L. Brunsveld, *Chem.Biol.*, 2009**, 16**, 702-711
2. T. Phan, H. D. Nguyen, H. Goksel, S. Mocklinghoff, and L. Brunsveld, *Chem.Commun.(Camb.)*, 2010**, 46**, 8207-8209
3. H. D. Nguyen, D. T. Dang, J. L. van Dongen, and L. Brunsveld, *Angew.Chem.Int.Ed Engl.*, 2010**, 49**, 895-898
4. S. Mocklinghoff, R. Rose, M. Carraz, A. Visser, C. Ottmann, and L. Brunsveld, *Chembiochem.*, 2010**, 11**, 2251-2254
5. B. Vaz, S. Mocklinghoff, S. Folkertsma, S. Lusher, V. J. de, and L. Brunsveld, *Chem.Commun.(Camb.)*, 2009 5377-5379
6. A. Tossi, L. Sandri, and A. Giangaspero, *Peptides*, 2002**, 27**, 416-417
7. M. R. Wilkins, E. Gasteiger, A. Bairoch, J. C. Sanchez, K. L. Williams, R. D. Appel, and D. F. Hochstrasser, *Methods Mol.Biol.*, 1999**, 112**, 531-552
8. C. Chang, J. D. Norris, H. Gron, L. A. Paige, P. T. Hamilton, D. J. Kenan, D. Fowlkes, and D. P. McDonnell, *Mol.Cell Biol.*, 1999**, 19**, 8226-8239